

caBIG:
Perspectives and Challenges
Albert Einstein Cancer Center
and Collaborating Institutions

For the AECC Informatics Group
(presented by Erwin Bottinger)

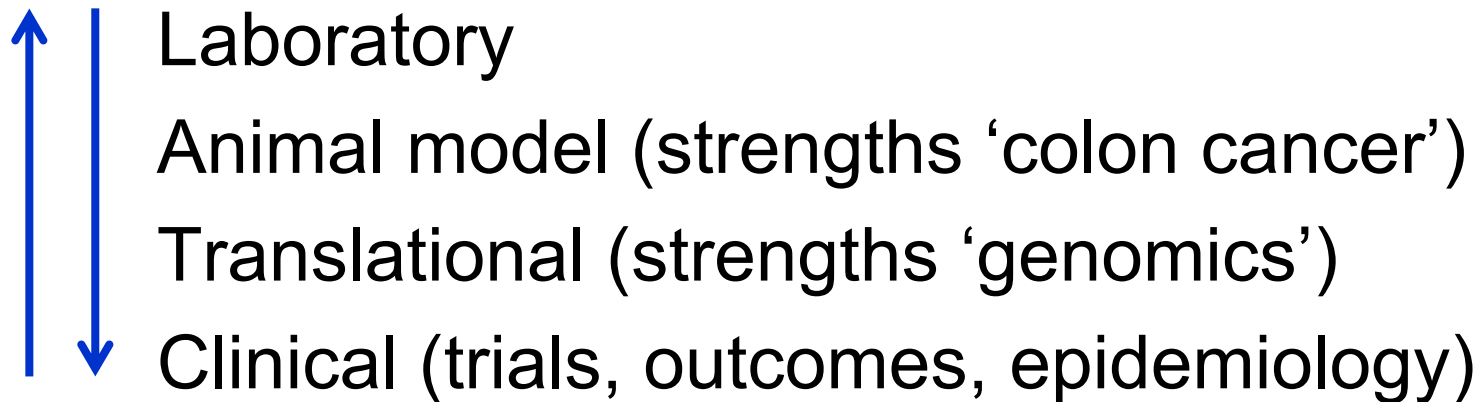
Collaborating Institutions:
Montefiore Medical Center
The Mount Sinai Medical Center

Perspectives

- Research Portfolio
 - Laboratory
 - Animal model (strengths ‘colon cancer’)
 - Translational (strengths ‘genomics’)
 - Clinical (trials, outcomes, epidemiology)
- Informatics Portfolio
 - Bioinformatics
 - Medical Informatics

Challenges

- Vertical Data Integration ?



- Bio-Medical-Informatics ?



Example

AECC Colon Cancer Program (Len Augenlicht)

- Scientific Theme: Interaction of Environmental and Genetic Factors
 - Genes and Loci
 - Nutrition and metabolism
 - Environment
 - Intervention and Outcome
- Mouse models of Colon Cancer
 - Mouse genetics
 - Functional genomics
 - Structural genomics
 - Biostatistics
- Cellular and Molecular Mechanisms:
 - Cells derived from tumor sites in animal models
 - Cells derived from human cancers

Example cont'd

AECC Colon Cancer Program (Len Augenlicht)

- Molecular Classification of Colon Cancers: A Clinical Trial
 - Gene expression database
 - SNP database
 - Clinical Trials Management System
- Comparative genomics in mouse, rat, human
 - Linking experimental data and public genetic data across species
 - Automated gene and pathway annotation pipeline

Example cont'd

AECC Cancer Epidemiology Program

(Tom Rohan, Robert Burk)

CANCER EPIDEMIOLOGY

Scientific Elements

Etiology

- HPV and cervical neoplasia
- Diet and cancer
- Genetic polymorphisms & cancer
- Gene discovery

Pathogenesis

- Cervical and breast cancer

Prevention

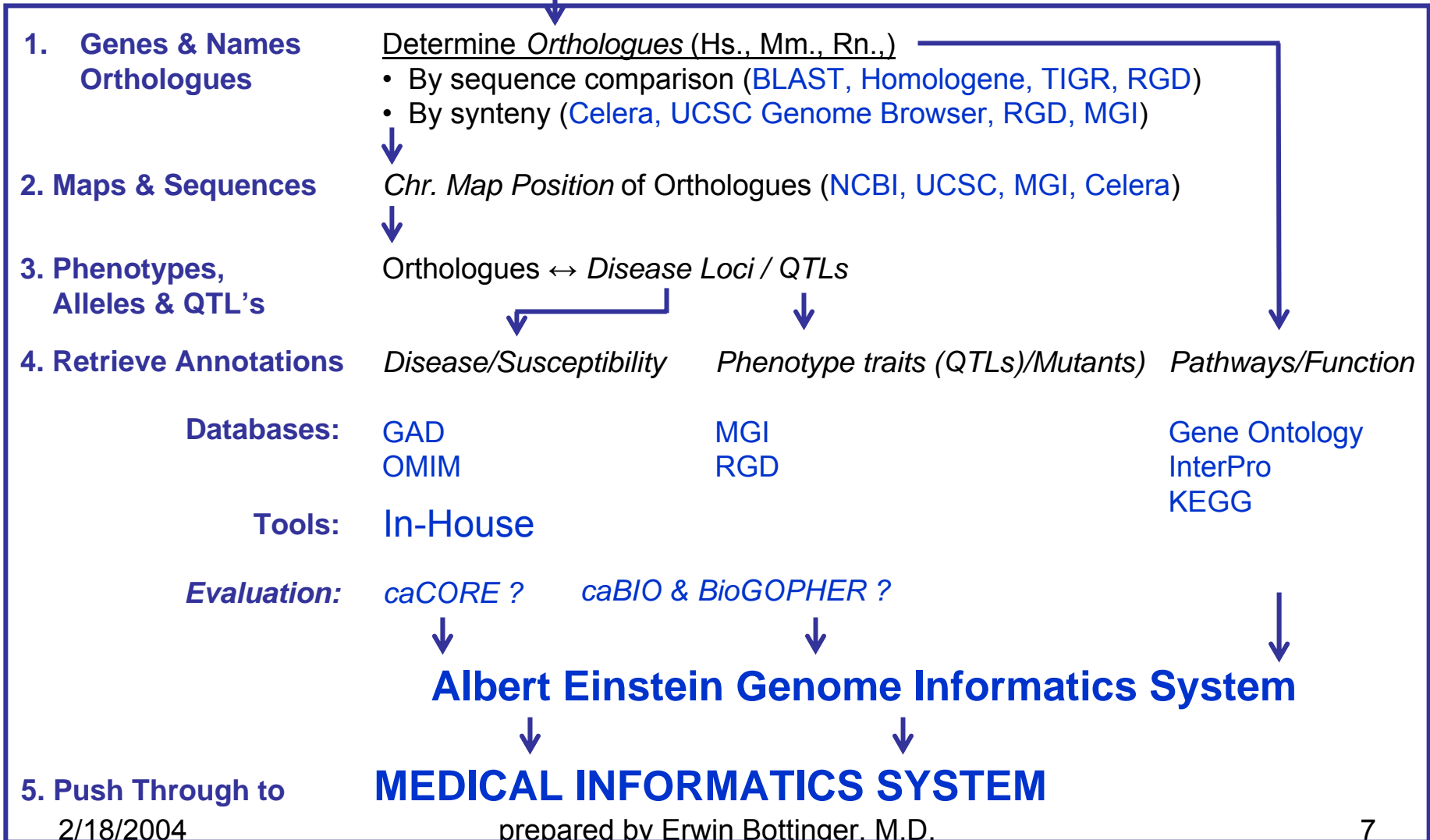
- Dietary
- Chemoprevention
- Behavioral

Bioinformatics Workspace: Comparative Genomics

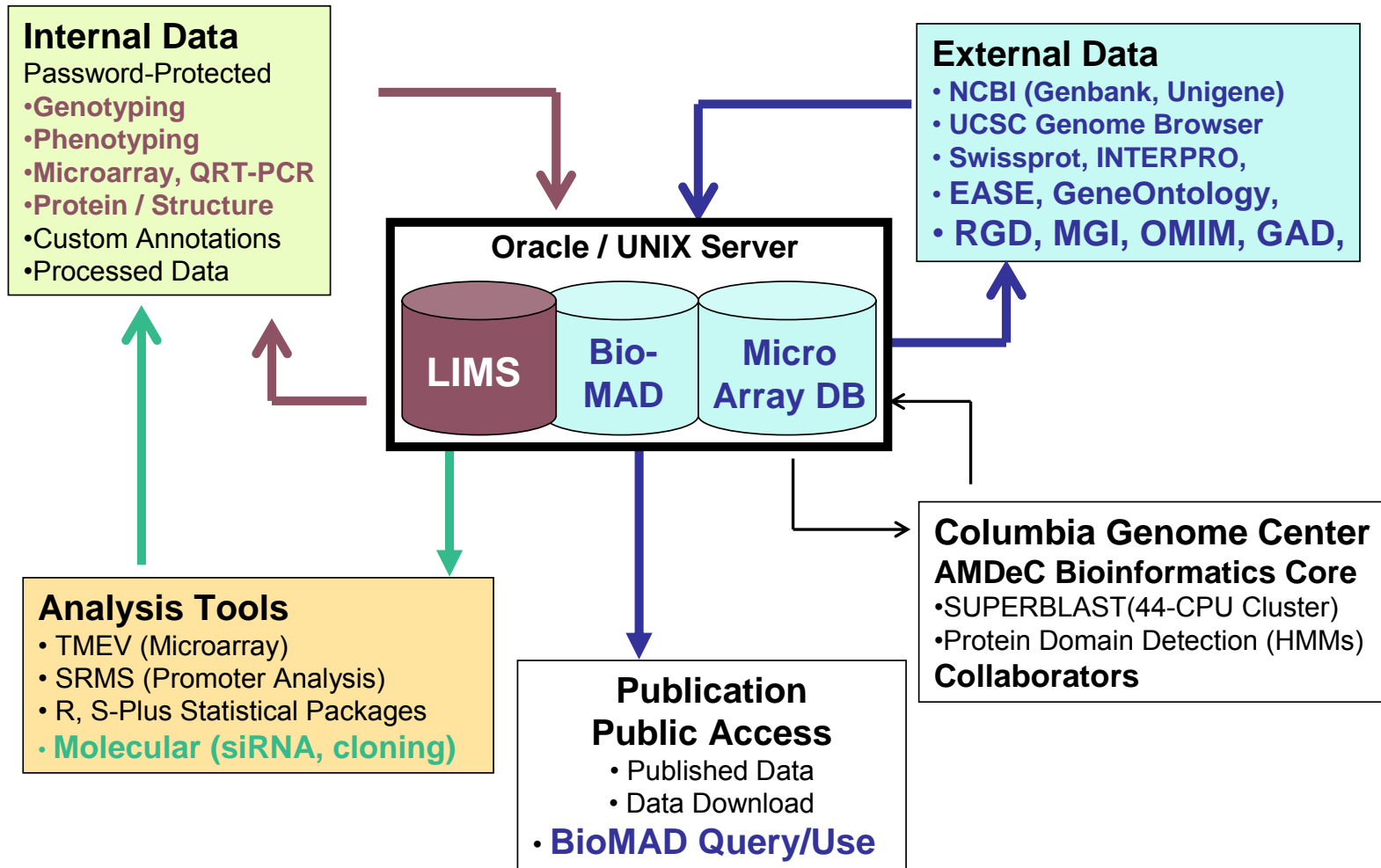
Gene Discovery and Biomarkers

Experimental Data Analysis

- *Genesets* w/ relevant expression profile (available and future)



DATA DOMAINS and DATA FLOW



Medical Informatics Resources - MSMC

Enterprise Data Repository (EDR)

- ICD-9, SNOMED, other vocabularies

Date Capture Start Date

Data Captured

1-Aug-00	Basic Demographic Information & Visit Summary for Hospital Patients
14-Aug-00	Radiology Inpatients and Outpatients Reports
17-Aug-00	Laboratory Inpatients Results and Micro Reports
18-Aug-00	Pathology Inpatients Reports
13-Oct-00	Laboratory Outpatient Results
24-Jan-01	Discharge Summary and Operative Reports
19-Apr-01	Pathology Outpatient Results

Sign-Out System Reports (SOSP)

- Clinical vocabulary / real-time

Goal: Integration of Bioinformatics and Medical Informatics Domains

What we will bring to caBIG :

- ‘Vertical Biomedical Informatics Environment’
 - Data environment and expertise in cancer research
 - Biology vocabs and DEs
 - Genomic vocabs and Des
 - Mouse model organism vocabs
 - Infrastructure for Implementation in cancer research and cancer care environment

What we hope to gain from caBIG :

- Workgroup support and applications for
 - Implementation and integration of clinical cancer domain vocabs and cancer biology vocabs
 - Implementation of data standards
 - Implementation of data exchange formats, i.e. XML
 - Networking